DOCUMENT RESUME

ED 199 047 SE 034 154

AUTHOR Freeman, Donald J.; And Others
TITLE The Fourth-Grade Mathematics Curriculum as Inferred From Textbooks and Tests. Research Series No. 82.
INSTITUTION Michigan State Univ., East Lansing. Inst. for Research on Teaching.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
REPORT NO MSU-IRT-RS-82
PUB DATE Jul 80
CONTRACT 400-76-0073
NOTE 32p.: Not available in hard copy due to marginal legibility of original document.
AVAILABLE FROM Institute for Research on Teaching, College of Education, 252 Erickson Hall, Michigan State University, East Lansing, MI 48824 ($2.50).

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Content Analysis; Course Content; *Educational Research; Elementary Education; *Elementary School Mathematics; *Grade 4; Mathematical Concepts; *Mathematics Curriculum; Mathematics Instruction; Test Items; *Tests; Textbook Content; *Textbooks
IDENTIFIERS *Mathematics Education Research

ABSTRACT

The results of one of the earliest studies in a series of investigations on teachers' selection of content for mathematics instruction suggested that teachers are responsive to content messages that are communicated through textbooks and standardizes tests. The content analyses that are summarized in this report were therefore undertaken in an attempt to identify similarities and differences in the "implied curricula" of these two sources of influence. All items in the student exercise portions of three widely used fourth-grade mathematics textbooks and all mathematics items in five commonly administered standardized tests were classified using a three-dimensional taxonomy. The results suggest that only a limited number of specific topics are consistently emphasized in textbooks and tests of fourth-grade mathematics. Further, the match in content is better for some combinations of textbooks and tests than for others. However, even the best matched textbooks and tests fail to provide a consistent description of the fourth-grade mathematics curriculum. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Lawrence Lezotte

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

SCOPE OF INTEREST NOTICE

The ERIC Facility has assigned this document for processing to: SE TM

In our judgement, this document is also of interest to the clearinghouses noted to the right. Indexing should reflect their special points of view.

ED199047

Institute for Research on Teaching

College of Education — Michigan State University

SE 034 154

Research Series No. 82

THE FOURTH-GRADE MATHEMATICS CURRICULUM AS INFERRED FROM TEXTBOOKS AND TESTS

Donald J. Freeman, Therese M. Kuhs,
Andrew C. Porter, Lucy B. Knappen,
Robert E. Floden, William H. Schmidt,
and John R. Schwille

Published By

The Institute for Research on Teaching
252 Erickson Hall
Michigan State University
East Lansing, Michigan 48824

July, 1980

This work is sponsored in part by the Institute for Research on Teaching, College of Education, Michigan State University. The Institute for Research on Teaching is funded primarily by the Program for Teaching and Instruction of the National Institute of Education, United States Department of Health, Education, and Welfare. The opinions expressed in this publication do not necessarily reflect the position, policy, or endorsement of the National Institute of Education. (Contract No. 400-76-0073)

JAN 2 6 1981

Institute for Research on Teaching

Teachers' thoughts and decisions are the focus of studies currently under way at Michigan State University's Institute for Research on Teaching (IRT). The IRT was founded in April 1976 with a $3.6 million grant from the National Institute of Education. That grant has since been renewed, extending IRT's work through September 1981. Funding is also received from other agencies and foundations. The Institute has major projects investigating teacher decision-making, including studies of reading diagnosis and remediation, classroom management strategies, instruction in the areas of language arts, reading, and mathematics, teacher education, teacher planning, effects of external pressures on teachers' decisions, socio-cultural factors, and teachers' perceptions of student affect. Researchers from many different disciplines cooperate in IRT research. In addition, public school teachers work at IRT as half-time collaborators in research, helping to design and plan studies, collect data, and analyze results. The Institute publishes research reports, conference proceedings, occasional papers, and a free quarterly newsletter for practitioners. For more information or to be placed on the IRT mailing list please write to : The IRT Editor, 252 Erickson, MSU, East Lansing, Michigan 48824.

Co-Directors: Judith E. Lanier and Lee S. Shulman

Associate Directors: Lawrence W. Lezotte and Andrew C. Porter

Editorial Staff:

Lawrence W. Lezotte, coordinator of Communications/Dissemination
Janet Flegg, IRT editor

Abstract

The results of one of the earliest studies in a series of investigations on teachers' selection of content for mathematics instruction suggested that teachers are responsive to content messages that are communicated through textbooks and standardized tests. The content analyses that are summarized in this report were therefore undertaken in an attempt to identify similarities and differences in the "implied curricula" of these two sources of influence. All items in the student exercise portions of three widely used fourth-grade mathematics textbooks and all mathematics items in five commonly administered standardized tests were classified using a three-dimensional taxonomy. The results suggest that only a limited number of specific topics are consistently emphasized in textbooks and tests of fourth-grade mathematics. Further, the match in content is better for some combinations of textbooks and tests than for others. However, even the best matched textbooks and tests fail to provide a consistent description of the fourth-grade mathematics curriculum.

THE FOURTH-GRADE MATHEMATICS CURRICULUM AS INFERRED FROM TEXTBOOKS AND TESTS[1]

Donald J. Freeman, Therese M. Kuhs, Andrew C. Porter,
Lucy B. Knappen, Robert E. Floden, William H. Schmidt,
and John R. Schwille[2]

Recent research suggests that there is a direct relationship between what is taught and what is learned (Walker & Schaffarzick, 1974; Schmidt, Note 1; Wilson, Cohen, & Begel, Note 2). Evidence also suggests that there is considerable variation in what is taught across different classrooms (Knappen & Floden, Note 3). Because variation in content covered will result in variation in student achievement, it is important to examine what causes content variation in the first place.

The work of the IRT Content Determinants Project is based on a conceptual model which suggests that various factors may directly or indirectly influence a teacher's selection of the content of instruction (See Porter, Schwille, Floden, Freeman, Knappen, Kuhs, & Schmidt, Note 4). When a district decides to adopt a single textbook or develops a list of instructional objectives, there is an implicit expectation that teachers will move toward the curriculum that is implied by each of these sources of influence. One of our earliest studies suggests that teachers are responsive to these and other pressures to teach certain content (Floden, Porter, Schmidt, Freeman, & Schwille, Note 5). It is therefore important to examine the

---

[1]Paper presented at the annual meeting of the American Educational Research Association, Boston, 1980.

[2]All the authors are members of the Content Determinants Project research team. Freeman, Floden, Schmidt, and Schwille are senior researchers; Kuhs is an intern; Porter is project coordinator; and Knappen is a teacher collaborator. The authors would like to thank Sister Chrisanne Weisbeck and Lenore Smith, who served as raters in the analysis reported here, and Suwatana Sookpokakit, who provided technical assistance.

"implied curriculum" of each source of influence to determine the content message that is communicated to the teacher through that factor.

Textbooks and standardized tests may be particularly strong determinants of the content of classroom instruction. The influence of textbooks on what is taught is generally recognized and supported by surveys that report that large numbers of classroom teachers follow the curricula of textbooks closely (National Advisory Committee on Mathematics Education, Note 6). The trend toward minimal competency testing, on the other hand, implies that teachers should teach what is tested. Thus standardized tests are also a plausible source of influence on the content of classroom instruction. In fact, there is some evidence to suggest that teachers are especially sensitive to this source of pressure (Floden et al., Note 5).

Recognition that textbooks and tests may influence content selection motivated our earlier investigations of the implied curricula of these two sources. Because the Content Determinants Project focuses on the fourth-grade mathematics curriculum, these investigations examined textbooks and tests in this content area. Our preliminary analyses of fourth-grade mathematics textbooks published by Addison-Wesley, Houghton-Mifflin, and Scott-Foresman suggested that there are important similarities and differences in content covered in these three books (Porter, Kuhs, & Freeman, Note 7). All three books reflect the current back to basics and minimal competency trends in their common emphasis on computational skills. However, more than half of the specific topics covered across the three books are unique to the book in which they appear.

Our preliminary analyses of five tests of fourth-grade mathematics, published by Harcourt Brace Javonovich, McGraw-Hill, and Houghton Mifflin, suggest that there are comparable similarities and differences in content treatment across tests (Freeman, Kuhs, Knappen, & Porter, in press).

Whereas the five tests are reasonably consistent in their treatment of whole number computational skills, there are striking differences in their treatment of such general topics as applications involving graphs or tables and number sentences. Further, there are only nine specific topics common to all five tests.

These preliminary analyses prompted two important questions that have served as the focus of our investigation. First, given that certain topics are emphasized in all three textbooks, it is important to determine how these particular topics are covered by tests. In other words, is the implied curriculum that is clearly communicated across all textbooks congruent with that which is communicated across all tests? Second, because content treatment varies among textbooks and tests, it is important to determine if the content of a particular textbook is better reflected in one test than another. In other words, is the match between the implied curricula communicated through textbooks and tests better for some combinations of textbooks and tests than for others?

Procedure

Selection of Textbooks and Tests

Four widely used standardized tests and three commonly adopted textbooks were selected for these comparisons. The tests are the following:

MAT -- Metropolitan Achievement Tests (Elementary Level/Grades 3.5-4.9), Harcourt Brace Javonovich, Inc., 1978.

SAT -- Stanford Achievement Test (Intermediate Level/Grades 4.5-5.6), Harcourt Javonovich, Inc., 1973.

CTBS -- Comprehensive Tests of Basic Skills (Level I/Grades 2.5-4.9 and Level II/Grades 4.5-6.9), McGraw-Hill, 1976.

IOWA -- Iowa Test of Basic Skills (Level 10/Grade 4), Houghton Mifflin, 1978.

As the above descriptions suggest, the administration of each of these tests at the end of the fourth-grade year would serve as an appropriate

measure of student achievement in fourth-grade mathematics. However, a school district that uses the CTBS for this purpose is faced with a decision between using Level I or Level II. Thus both versions have been included in the analysis.

The three textbooks of fourth-grade mathematics selected for the analysis are as follows:

Mathematics in Our World, Addison-Wesley Publishing Co., 1978.

Mathematics, Houghton-Mifflin Co., 1978

Mathematics Around Us, Scott-Foresman and Co., 1978.

The four tests are similar in that each contains sets of items that deal with mathematical concepts, computations, and applications. The three textbooks are also similar in organization. Lessons in all three books contain two distinct components: instructional activities directed by the teacher and practice exercises assigned to students. Whereas our content analyses of tests considered all items in each subtest, our analyses of textbooks were limited to items in the student exercise portions of each lesson. The decision to limit our analyses of topics to these portions was based on practical considerations as well as the results of a survey conducted by the National Advisory Committee on Mathematics Education (Note 6), which indicated that a majority of teachers rely primarily on the student exercise portion of each lesson.

The Taxonomy on Which the Analyses were Based

The content analyses of textbooks and tests were guided by a classification manual that describes a three-dimensional taxonomy developed for this purpose (Kuhs, Schmidt, Porter, Floden, Freeman, & Schwille, Note 8). The three dimensions of the taxonomy describe the general intent of the item (e.g., conceptual understanding or application), the nature of material presented to students (e.g., measurement

or decimals), and the operation the student must perform (e.g., estimate or multiply). One of the most significant features of the taxonomy is its flexibility in describing content at different levels of detail. Intersections of conditions along all three dimensions describe specific topics that teachers might elect to cover in the classroom (e.g., solving story problems that require multiplication of single digit whole numbers). At the same time, descriptions that correspond to the marginals of the taxonomy represent general topics that might be included in the mathematics curriculum (e.g., fractions, division with remainders, story problems).

Level of Interrater Agreement

Content analyses for both textbooks and tests were done independently by two raters. All items on the standardized tests and every item in the student exercise portion of each textbook lesson were described along all three dimensions of the taxonomy. The results for a given pair of raters were then compared, and discrepancies were resolved by a third rater.

Although there was some variance among the five raters who participated in the study, interrater agreements were generally very high. Interrater agreements across the four standardized tests were consistently above .98. The corresponding figures for the Houghton-Mifflin text exceeded .97 for both marginal and cell-level totals. A complete description of interrater reliabilities will be provided in a forthcoming technical report.

Content Measures

Although a single measure, percent of items dealing with a given content category provides a satisfactory picture of content treatment in standardized tests, there are at least three meaningful measures of content covered in textbooks: percent of items, percent of lessons, and proportion of items within each lesson that deal with a given topic. Because

the percent of lessons measure does not facilitate meaningful comparisons of content covered in textbooks and tests, it is not considered in this report.

Item percents for both textbooks and tests were determined by dividing the sum of items that deal with a given content category by the total number of items in the textbook or test. Adjusted lesson percents were determined by dividing the sum of the proportion of items in each lesson that deal with a given topic by the total number of lessons in the book. The calculations on which these two measures were based are illustrated in Figure 1.

Lesson #	Number of Items Dealing with Fractions	Total # of Items	Proportion of Items Dealing with Fractions
1	0	40	0%
2	20	40	50%
3	20	20	100%
4	2	20	10%

$$\text{Item percent} = \frac{0 + 20 + 20 + 2}{40 + 40 + 20 + 20} = 35\%$$

$$\text{Adjusted lesson percent} = \frac{0 + 50 + 100 + 10}{4} = 40\%$$

Figure 1. An illustration of the derivation of item percent and adjusted lesson percent measures.

Although some may question the need for adjusted lesson percents, the decision to include this measure was based on the observation that items dealing with certain content categories (e.g., story problems) are characteristically longer and more complex than items dealing with other content categories (e.g., multiplication facts). Thus student exercises that focus on the more complex skills tend to contain fewer items, and totals for these categories are underrepresented in the percent of items

measure.

Results

Congruity in the Core Curricula of Textbooks and Tests

Imagine that a school district mandates the use of one of the three textbooks considered in this investigation and administers one of the four tests at the end of the fourth-grade year. If a teacher examines the textbook and test to get some sense of what should be taught in fourth-grade mathematics, what content message, if any, will be clearly communicated by both the textbook and test? In order to address this question we must determine the extent to which the implied curriculum communicated across all textbooks is congruent with what is communicated across all tests.

We refer to the set of topics that is emphasized in all three textbooks as the core curriculum suggested by textbooks. A specific topic is arbitrarily included in the core curriculum if it is common to all three books and represents at least 0.5% of the adjusted lesson percents for each book. The content emphasis criterion is functionally equal to approximately .7 of one lesson in the Addison-Wesley book, .8 of one lesson in the Houghton-Mifflin book,and one full lesson in the Scott-Foresman book. This variation is due to differences in the total number of lessons in each book. Because core topics are emphasized in all three books, it is likely that teachers will receive the message to teach them regardless of the book they are using.

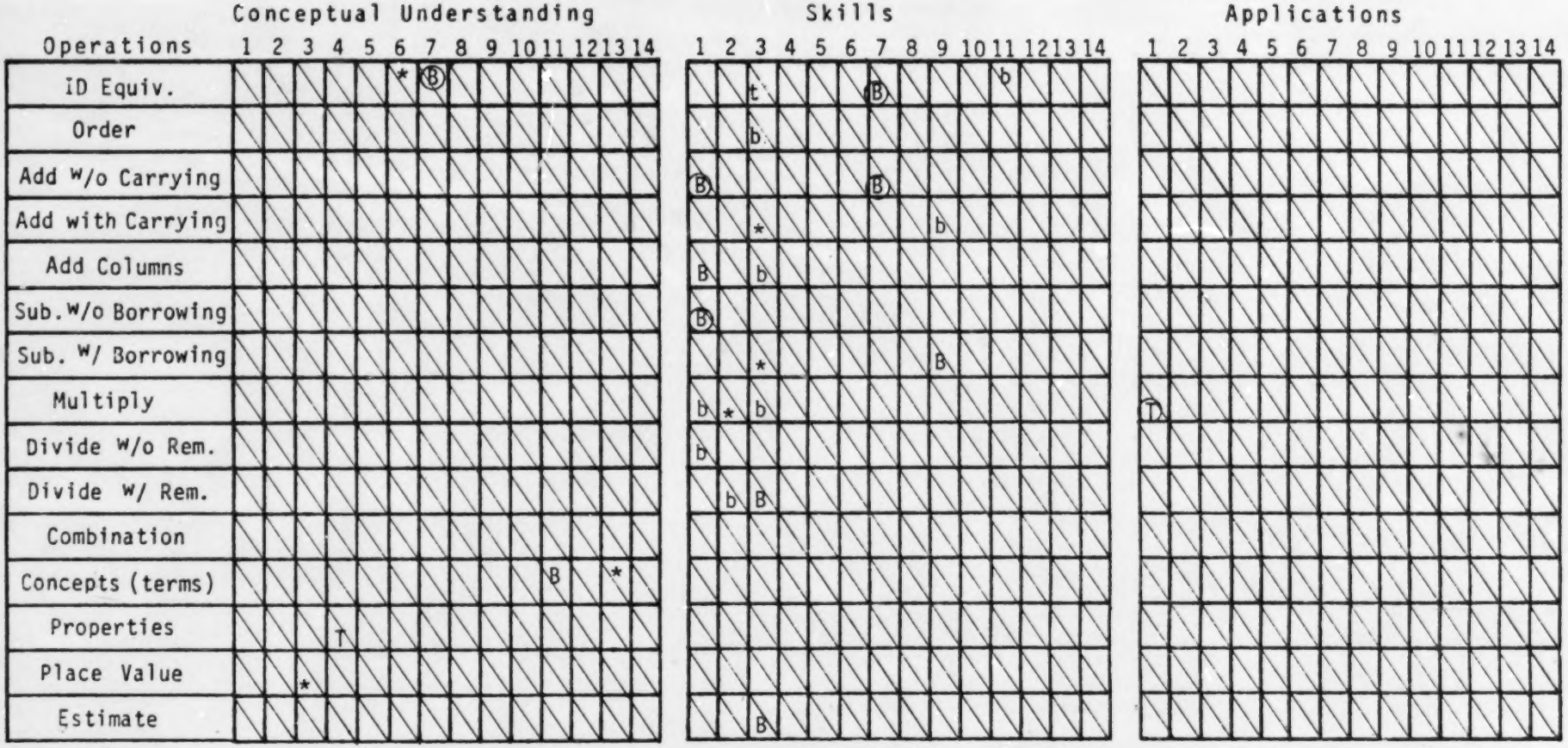
We refer to the set of topics covered across the five tests as the core curriculum suggested by standardized tests. A specific topic is arbitrarily included in this core curriculum if there is at least one item that deals with the topic on the MAT, SAT, IOWA, and either the CTBS I or CTBS II. The fact that the two forms of the CTBS are designed for use at

different grade levels, yet overlap at the end of the fourth grade, necessitates the either/or condition in this decision rule. The fact that one item is equivalent to approximately .2 of a grade equivalent near the middle of the norm distribution for subtests suggests that a single item is a reasonable standard for content emphasis. Because there is at least one item that deals with these topics on each of the tests, it is likely that teachers will receive the message to teach core topics when they critically examine tests, regardless of which test they are analyzing.

The 27 specific content areas that satisfy one or both definitions for core topics are portrayed in Figure 2. As this figure suggests, most of these topics are described by the skills dimension of the taxonomy; only one topic involves applications. Further, the intersection of core curricula for textbooks and tests includes only six specific topics, three of which focus on conceptual understanding.

Figure 3 provides descriptions of the 27 core topics. The five core topics unique to textbooks are described at the beginning of this list; the single topic that is unique to tests is the final entry on the list. As the format of this figure suggests, there is at least some overlap in how each of the other 23 topics is treated in textbooks and tests. Figure 3 also describes the adjusted lesson percents across the three textbooks and the percent of items that focus on a given topic across the five tests. These statistics provide an index of content emphasis.

The information that is summarized in Figure 3 provides graphic evidence that the fourth-grade mathematics curriculum suggested by textbooks may be characterized by its concern with skills. Of the 24 core topics, 19 are in the skills area. Multiplication and division of whole numbers and basic facts account for approximately 25% of the lessons in all three books. At the other extreme, there are no core textbook topics that



Ⓑ - All books, no tests
B - All books, one test
b - All books, 2 or 3 tests
* - All books, all tests

Ⓣ - All tests, no books
T - All tests, one book
t - All tests, two books

W/out ▨ with pictures

Nature of Material

1. sing. dig./basic facts
2. sing. & mult. digit
3. multiple digit
4. no. sen./phrase
5. alg. sen./phrase
6. sing./like frac.
7. unlike frac.
8. mixed no.
9. decimals
10. percents
11. measurement
12. essn. units of measurement
13. geometry
14. other

Figure 2. Core curricula of textbooks and tests.

Topic Description	Adjusted Lesson Percents			Percent of Items			
	A.W.	H.M.	S.F.	CTBSI	CTBSII	IOWA	MAT	SAT
# Lessons	(145)	(167)	(194)	# Items (98)	(98)	(103)	(50)	(112)
I. Topics emphasized in all three books but none of the tests								
Pictorial models of = fractions	0.69	0.65	2.20	0	0	0	0	0
Basic Addition facts	0.98	2.56	0.99	0	0	0	0	0
Basic Subtraction facts	1.16	1.60	1.04	0	0	0	0	0
Find equivalent fractions	1.77	1.81	0.59	0	0	0	0	0
Add fractions (unlike denominators)	0.85	0.87	0.51	0	0	0	0	0
II. Topics emphasized in all three books but only one test								
Pictorial model-measurement term	1.28	0.59	0.55	0	0	1.0	0	0
Column addition (single digit #)	1.00	1.07	0.93	2.0	0	0	0	0
Division w/remainder (multiple digit #)	2.12	1.76	2.42	0	1.0	0	0	0
Estimating or rounding off	1.31	1.19	1.36	0	0	1.0	0	0
Subtract w/borrowing (decimals)	2.12	0.60	0.96	0	3.1	0	0	0
III. Topics emphasized in all three books and on two or three tests								
Basic Multiplication facts	5.62	7.31	3.91	3.1	0	2.9	4.0	0
Basic Division facts	3.81	5.08	3.14	9.2	1.0	1.9	4.0	0
Division w/remainder (single digit divisor)	2.32	4.10	2.24	0	1.0	1.0	0	0
Ordering Multiple digit #	0.52	1.44	0.77	1.0	0	0	6.0	0
Column Addition (multiple digit #)	1.55	0.54	0.52	4.1	4.1	3.9	0	2.7
Multiplication (multiple digit #)	4.61	3.56	3.20	0	2.0	2.9	0	2.7
Addition w/carrying (decimals)	1.83	0.73	0.61	1.0	0	0	0	0.9
Read Measurement Instruments	2.21	2.20	2.86	1.0	1.0	0	2.0	0
IV. Core topics in textbooks and tests								
*Pictorial Models of a fraction	1.89	1.14	2.21	1.0	2.0	1.0	2.0	0.9
*Geometric terms (w/pictures)	4.33	4.93	7.65	5.1	4.1	1.9	2.0	1.8
Place value (multiple digit #)	1.58	1.45	2.85	2.0	0	2.9	4.0	3.6
*Addition w/carrying (multiple digit #)	1.61	1.95	1.57	4.1	1.0	4.9	2.0	0.9
*Subtraction w/borrowing (multiple digit #)	3.91	2.16	1.88	2.0	4.1	6.8	4.0	3.6
*Multiplication (single x multiple digit #)	3.50	7.32	3.56	9.2	4.1	5.8	4.0	2.7
V. Topics included on all tests but not emphasized in all books								
Reading large numbers	1.54	0.16	1.73	2.0	1.0	1.9	2.0	0.9
# sentences involving properties	0	0	1.11	2.0	0	1.9	2.0	2.7
Story problems (basic multiplication facts)	0.37	0.31	0.13	1.0	0	1.0	2.0	0.9

* Topics covered on both forms of CTBS.

Figure 3. A description of the core curriculum of textbooks and tests.

involve applications. Even if the emphasis criterion in the definition of core topics were reduced to the equivalent of 25% of one lesson in each book, only six core topics would emerge. These may be described as perimeter calculation, story problems involving division facts, and story problems involving one of the following forms of multiplication: multiplication facts, one-digit multipliers, multiple-digit multipliers, or quantities expressed in measurement units.

Although the core curriculum suggested by standardized tests consists of only nine topics, these topics are more evenly distributed across the three general dimensions of the taxonomy than is true for books. Four core test topics involve skills, four focus on conceptual understanding, and one involves applications. Given that the range of the core curriculum suggested by tests is restricted to nine specific topics, one might expect that all of these topics would be emphasized in fourth-grade mathematics textbooks.

However, the intersection of core topics suggested by both textbooks and tests consists of only six specific topics, three of which involve skills and three of which focus on conceptual understanding. In other words, there are only six specific areas within the fourth-grade mathematics curriculum that are consistently emphasized across the three textbooks and five tests considered in this analysis.

Thus a teacher who examines a textbook or a standardized test in an attempt to get some sense of what to teach in fourth-grade mathematics should receive the message to teach geometric terms, the concept of a fraction as represented by pictorial models, place value concepts, addition of multiple-digit numbers with carrying, subtraction of comparable numbers with borrowing, and multiplication with one-digit multipliers and multiple digit multiplicands. Any other content message the teacher might

receive will vary as a function of the particular textbook or test the teacher has examined.

Some might argue that the number of topics we have identified as common to all textbooks and all standardized tests has been severely restricted by the analysis procedures we used. Some, for example, might contend that our definitions of core topics are too stringent; others might argue that our content categories are too specific. Let us therefore consider both of these concerns.

As noted earlier, the criteria we used in identifying core topics were arbitrarily determined. Therefore, imagine that these criteria had been relaxed such that any topic that is represented by at least one item in each of the three textbooks and appears on at least three of the five standardized tests would be classified as a topic that is commonly covered in both textbooks and tests. Even with these relaxed criteria, only 16 specific topics would be classified as common to textbooks and tests (see Levels III, IV, and V of Figure 3).[3] Twelve of these topics involve skills, three focus on conceptual understanding, and one involves applications.

In brief, regardless of what standards are used in identifying specific topics that are consistently covered in both textbooks and tests, the list of common topics will be very short. In other words, most specific topics that one might attribute to the fourth-grade mathematics curriculum are covered in some textbooks and/or standardized tests, but do not appear in others.

The second factor that may have limited the number of topics that we identified as common to all textbooks and all standardized tests is

---

[3]The Addison-Wesley and Houghton-Mifflin texts do not contain any items that focus on properties expressed in number sentences.

the level of detail of the content categories we described. There is little doubt that the level of congruity of content covered in textbooks and tests is a function of the level of detail with which content is described; the greater the specificity, the lower the level of congruity. It is therefore reasonable to ask what constitutes an appropriate level of detail.

Our approach to this issue is functionally tied to the basic purpose of our research. Given our desire to delineate factors that influence content selection, we have deliberately attempted to develop a taxonomy that will provide accurate descriptions of mathematics topics that teachers may elect to cover or not to cover in their classrooms. We have also undertaken a complex investigation of the taxonomy's ability to provide content descriptions of this type (Schwille, Note 9). Although still in the early stages of development, impressions we have gained from this investigation suggest that teachers describe content decisions involving conceptual understanding or skills in as much or more detail than that suggested by cells in our taxonomy. Thus we have at least some reason to believe that the content categories we have considered in our analysis are generally consistent with teachers' descriptions of content decisions involving conceptual understanding or skills.

Unfortunately, our impressions regarding the characteristic level of detail of teachers' descriptions of application topics are less clear. At this point, we must therefore consider the possibility that descriptions of topics involving applications provided by the cell level of our taxonomy are more detailed than those that characterize a teacher's selection of topics to cover in the classroom.

By aggregating across cells, it is possible to identify general topics that might be covered in books and tests. If we ignore distinctions made

in the nature of material or operations dimensions of our taxonomy, for example, it is possible to show that all textbooks and tests do deal with application problems that require some form of multiplication, addition, subtraction, division, or some combination of two or more of the basic operations. However, even at this level of aggregation, there are distinct differences in the coverage of application topics in textbooks and tests. For example, some, but not all, books and tests emphasize application problems that require some form of estimation. Further, there is little or no consistency in how textbooks and tests deal with general categories of application problems in which information needed to solve the problem is conveyed in graphs or charts. In brief, regardless of how we aggregated across cells within the application dimension of the taxonomy, it was possible to identify only five conceptually meaningful content categories that were consistently emphasized across the three books and four tests.

Textbook/Test Comparisons

Imagine that a school district mandates the use of one of the three textbooks considered in this investigation and administers one of the four tests at the end of the fourth-grade year. To what extent will the level of consistency in content messages communicated to the teacher through these two sources vary as a function of the particular textbook and test used in the district?

There are several ways in which this question might be addressed. These analyses are suggested by the following questions: What proportion of topics covered on a given exam are also covered in each book? What

proportion of topics addressed in a given book are also covered in each test? Finally, what is the correlation between measures of content emphasis for each topic covered in the book and/or exam?

Table 1 describes the percent of tested topics covered in each textbook. The three columns labeled "T" describe the percent of specific topics on the exam that serve as the focus of at least one item in the student exercises in each book. These values range from a low of 54.2% for the SAT and Addison-Wesley text to a high of 74.4% for the MAT and both the Houghton-Mifflin and Scott-Foresman texts. Thus in even the best matched textbooks and tests, less than 75% of the specific topics that appear on the exam are considered in the book. The column labeled "T^1" describes the percent of test topics that represent at least 0.5% of the adjusted lesson percents for each book. These figures reflect the percent of test topics that are emphasized in the book, that is, topics that account for the equivalent of at least 80% of the items in one lesson. These values range from a low of 19.4% for the SAT and Houghton-Mifflin text to a high of 41.0% for the MAT and Scott-Foresman text. In other words, the proportion of topics that are presented on an exam that have received more than cursory treatment in the textbook may be as low as 20%.

A second way of thinking about the consistency of content covered on a given textbook and test is to ask what percent of the topics that are considered in a book are also covered on an exam. Table 2 describes the percent of textbook topics covered by each test. The rows labeled "T" represent the percent of all book topics covered in each test; the rows labeled "T^1" describe the percent of topics that are emphasized in the book (these account for approximately 80% or more of one lesson) and are also considered in the exam. The latter figures vary from a low of 21.8%

for the Addison-Wesley text and MAT to a high of 45.8% for the Houghton-Mifflin text and the Iowa Test of Basic Skills. In other words, for even the best matched textbook and test, less than 50% of the topics that are emphasized in the book serve as the focus of one or more items on the exam.

Table 1

Percent of Tested Topics Covered in Each Textbook*

	A-W		H-M		S-F	
	T (N=148)	T^1 (n=55)	T (N=167)	T^1 (n=48)	T (N=197)	T^1 (n=59)
MAT (n=39 topics)	61.5	30.8	74.4	38.5	74.4	41.0
SAT (n=72 topics)	54.2	22.2	52.8	19.4	62.5	26.4
IOWA (n=66 topics)	54.5	30.3	72.7	33.3	71.2	31.8
CTBS I (n=53 topics)	56.7	28.3	64.1	34.0	66.0	39.6
CTBS II (n=60 topics)	61.7	31.7	58.3	35.0	66.7	36.7

* T = Topics covered by at least one item in the book.

T^1 = Topics that represent at least 0.5% of the adjusted lesson percents for each book (approximately 4/5 of one lesson).

The figures in Tables 1 and 2 present an interesting contrast. Whereas the Scott-Foresman book covers the highest percent of topics tested on all but the Iowa exam (Table 1), this book also deals with a relatively high percent of topics that are not covered in the exams (Table 2). Thus it is difficult to determine which book provides the best match with the tests. A district that has already selected a standardized test and is in the process of selecting a book would probably find the figures in Table 1 most helpful; a district that has selected

a book and is searching for an exam might be better served by the figures in Table 2. Overall, it is clear from both tables that no one test is equally well suited for all textbooks and no one book is equally well suited for all tests.

Table 2

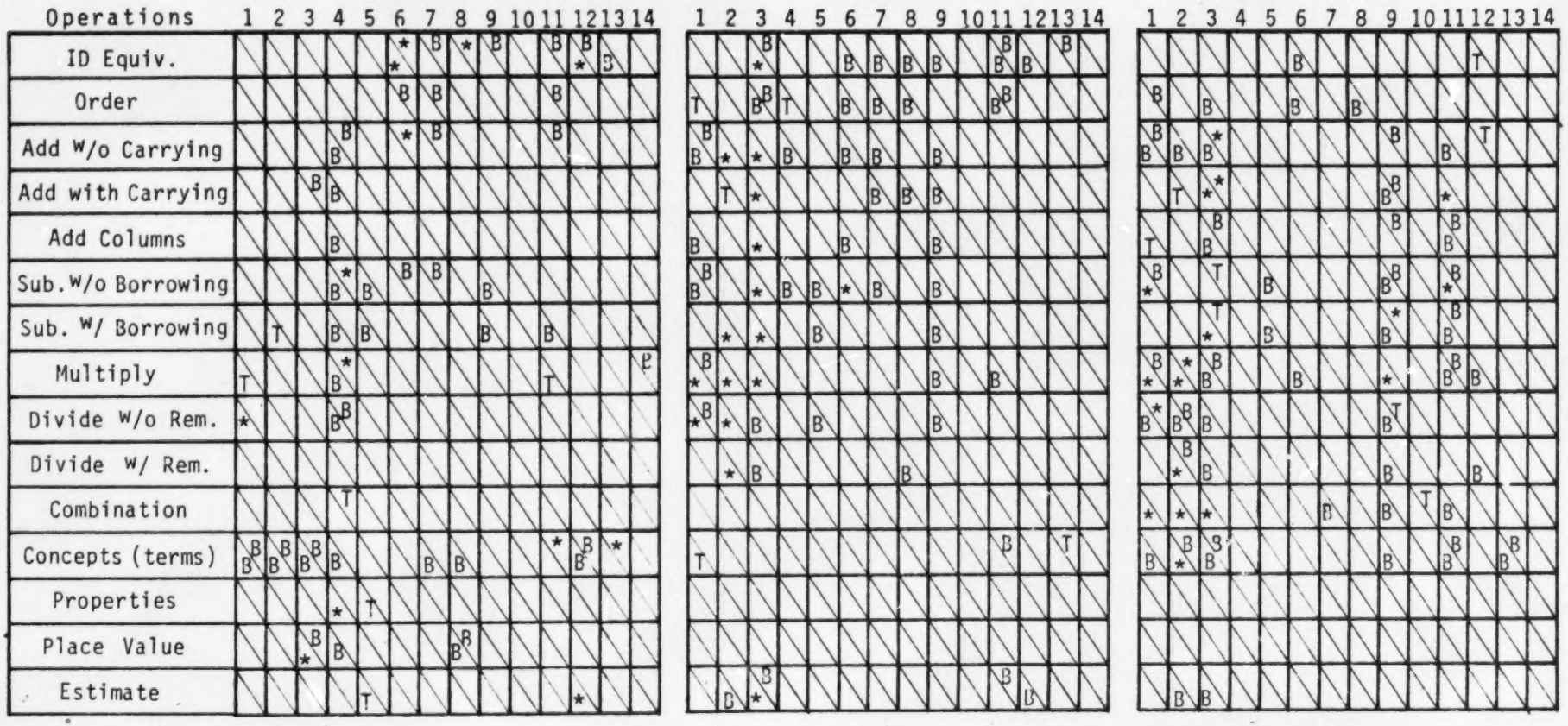
Percent of Textbook Topics Covered by Each Test*

		MAT (n=39)	SAT (n=72)	IOWA (n=66)	CTBS I (n=53)	CTBS II (n=60)
A-W	T (N=148 topics)	16.2	26.4	24.3	20.3	25.0
	T^1 (n=55 topics)	21.8	29.1	36.4	27.3	34.5
H-M	T (N=167 topics)	17.4	22.8	28.7	20.4	21.0
	T^1 (n=48 topics)	31.3	29.2	45.8	37.5	43.8
S-F	T (N=197 topics)	14.7	22.8	23.9	17.8	20.3
	T^1 (n=59 topics)	27.1	32.2	35.6	35.6	37.3

* T = Topics covered by at least one item in the book

T^1 = Topics that represent at least 0.5% of the adjusted lesson percents for each book (approximately 4/5 of one lesson)

Figure 4 shows the degree of consistency in content covered by the Scott-Foresman test and the Iowa Test of Basic Skills. As this figure suggests, the book and test were somewhat more consistent in their treatment of topics within the skills dimension of the taxonomy than their treatment of topics involving conceptual understanding or applications. This was generally true for all pairs of textbooks and tests. Figure 4 also provides graphic evidence that even when textbooks and tests are comparatively well matched, the majority of the topics in the total domain



* - in both book and test

B - in book only

T - in test only

W/out ▨ with pictures

Nature of Material

1. sing. dig./basic facts
2. sing. & mult. digit
3. multiple digit
4. no. sen./phrase
5. alg. sen./phrase
6. sing./like frac.
7. unlike frac.
8. mixed no.
9. decimals
10. percents
11. measurement
12. essn. units of measurement
13. geometry
14. other

Figure 4. Comparison of specific topics covered in the Scott Foresman text and the Iowa Test of Basic Skills.

of content covered in one or both sources are considered in one source but not the other. In other words, even the best matched textbooks and tests fail to provide a consistent description of the fourth-grade mathematics curriculum.

A third way in which the content covered in a given textbook and test may be compared is to determine correlations between measures of content emphasis for both sources. Table 3 presents textbook test correlations based on the percent of items in a given textbook and test that focused on each specific topic within the total domain of content covered in the book and/or test. Those topics that were not considered in either source were not included in this analysis.

As Table 3 suggests, content emphasis correlations ranged from a low of .12 for the SAT and Addison-Wesley text to a high of .55 for the CTBS I and Scott-Foresman text. The CTBS I provided the best match in content emphasis across all three books; the SAT provided the worst overall match.

The data summarized in Tables 1-3 provide convincing evidence that the match in content covered is better for some combinations of textbooks and tests than for others. Thus the level of congruity in content messages communicated to the teacher will vary dramatically as a function of the particular textbook and test the teacher has been asked to use.

Table 3

Test/Textbook Correlations (Based on Percent of Items for Each Specific Topic Covered in the Test and/or Book)

	MAT	SAT	IOWA	CTBS I	CTBS II
A-W	.28	.12	.44	.45	.22
H-M	.34	.17	.44	.54	.30
S-F	.39	.15	.46	.55	.28

Issues Suggested by this Analysis

Textbooks and standardized tests represent public statements of the elementary-school mathematics curriculum. Books may be viewed as statements of what should be taught, while tests represent descriptions of what might be learned. Given the public nature of these documents and the fact that tests are designed to measure that which has been learned, it is reasonable to expect a clear match in content covered in textbooks and tests of fourth-grade mathematics. However, the data described in this report provide convincing evidence for the following statements.

1. Only a limited number of specific topics are consistently emphasized in textbooks and tests of fourth-grade mathematics.

2. Even the best-matched textbooks and tests fail to provide a consistent description of the fourth-grade mathematics curriculum.

In short, there are striking differences in the implied curricula of textbooks and tests of fourth-grade mathematics. Thus one must consider the impact of these differences on a diverse array of significant educational issues. How should the lack of congruity in content covered in textbooks and tests influence the interpretation and application of standardized test scores? To what extent should an analysis of content covered guide the selection of textbooks and tests? To what extent are trends in student achievement at either the local or national level a function of changes in the level of congruity in content covered in textbooks and tests? Why are there significant differences in how textbooks and tests describe the elementary-school mathematics curriculum? How should teachers resolve differences in content messages communicated through textbooks and tests when deciding what to teach in mathematics?

We have dealt with the first two issues in other reports (See Porter, Schmidt, Floden, & Freeman, 1978; Freeman et al., in press). In our view,

there is a clear need to consider an analysis of content covered when interpreting standardized test scores or selecting textbooks and tests. The third issue is especially significant at this point in history when a great deal of public attention has been directed to the question of declining test scores. It is possible, for example, that this decline may be linked to a decline in the level of match between what is taught through textbooks and what is tested on standardized exams. In brief, there is a clear need for additional research that focuses on the content of classroom instruction as it relates to performance on standardized tests.

The issue of why textbooks and tests provide different descriptions of the mathematics curriculum is a complex one. There are a number of factors that are implicit in the design of standardized tests that suggest it would be unreasonable to expect a one-to-one correspondence between content covered in textbooks and tests for a single grade level. Standardized tests are designed to measure achievement at more than one grade level, represent samples and not the entire domain of content that is included in the curriculum, and are typically designed to measure transfer as well as achievement. Although each of these factors probably contributed to the differences in implied curricula of textbooks and tests identified in this investigation, it is unlikely that all of the differences can be simply ascribed to factors of this type. Rather, any effort to determine why textbooks and tests differ in how they describe the fourth-grade mathematics curriculum will have to consider a broad range of complex variables including political and economic factors that impact upon the development of textbooks and tests. Thus a comprehensive description of the origin of differences in the implied curricula of textbooks and tests is clearly beyond the scope and purpose of this report.

Regardless of origin, the fact remains that there are striking differences in how the curriculum is described by textbooks and tests. Thus one must consider how these differences impact upon a teacher's selection of the content of classroom instruction. Stated in simple terms, as teachers consider the question of what to teach, they must either resolve or ignore the conflicting content messages that are communicated through textbooks and tests.

Consider a fourth-grade teacher who is using the Houghton-Mifflin text and has been asked to administer the Stanford-Achievement Test at the end of the fourth-grade year. If this teacher carefully examines these two documents in an effort to get some sense of what to teach in fourth-grade mathematics, she will receive the message from the Houghton-Mifflin text to teach 167 specific topics. The Stanford test will suggest that she should teach 72 distinct topics, only 38 of which are included in the content message communicated through the book. Given that this teacher might also receive content messages through other sources such as district objectives or comments made by parents, her principal, or other teachers, it is readily apparent that some of these messages must be ignored. Given restrictions in the time available for mathematics instruction, it is simply not possible to provide adequate coverage of all of the topics she will be asked to teach.

But what topics should she ignore? There are no clear directives to guide these decisions. Thus it should come as no surprise that teachers develop unique patterns of responding to content messages and may choose to ignore even those messages (e.g., geometry terms) that are consistently communicated across various sources of influence (Knappen & Floden, Note 3).

This phenomenon will be especially troublesome for those who seek to bring about greater uniformity in the content of classroom instruction through various policy decisions such as mandated texts, district objectives, published test results, or some form of interpersonal persuasion. Given the inherent inconsistencies in content messages communicated across each of these sources of influence, striking differences in topics that teachers select for classroom instruction are virtually certain to occur even among those teachers who conscientiously seek to be responsive to the district's initiatives.

Reference Notes

1. Schmidt, W. A case study of school system objectives, a mandated test, and student achievement. Paper presented at the annual meeting of the American Educational Research Association, Boston, 1980.

2. Wilson, J.W., Cohen, L.S., & Begle, E.G. National Longitudinal Study of Mathematical Abilities Reports, Vols. 1-25. Stanford University: The Board of Trustees of Leland Stanford Junior University, 1968-72.

3. Knappen, L.B. & Floden, R. Who gets taught what when: Case studies of decisions about instructional content. Paper presented at the annual meeting of the American Educational Research Association, Boston, 1980.

4. Porter, A.C., Schwille, J.R., Floden, R.E., Freeman, D.J., Knappen, L.B., Kuhs, T.M., & Schmidt, W.H. Teacher autonomy and the control of content taught (Research Series No. 24). East Lansing, Michigan: Institute for Research on Teaching, Michigan State University, 1979.

5. Floden, R.E., Porter, A., Schmidt, W., Freeman, D., & Schwille, J. A policy-capturing study of teacher decisions about content. Paper presented at the meetings of the American Educational Research Association, Toronto, 1978. (Also available as IRT Research Series No. 74. East Lansing, Michigan: Institute for Research on Teaching, Michigan State University, 1980.)

6. National Advisory Committee on Mathematics Education. Overview and analysis of school mathematics: Grades K-12. Washington, D.C.: Conference Board on the Mathematical Sciences, 1975.

7. Porter, A.C., Kuhs, T.M., & Freeman, D.J. Is there a core curriculum in elementary school mathematics? Invited paper presented at the Textbook in American Education Conference co-sponsored by the Center for the Book in the Library of Congress and the U. S. National Institute of Education, May 2-3, 1979 (the paper will be published in a book of the proceedings).

8. Kuhs, T., Schmidt, W., Porter, A., Floden, R., Freeman, D., & Schwille, J. A Taxonomy for classifying elementary school mathematics content (Research Series No. 4). East Lansing, Michigan: Institute for Research on Teaching, Michigan State University, 1979.

9. Schwille, J. The content of elementary school mathematics from the perspective of teachers. Paper presented at the annual meeting of the American Educational Research Association, Boston, 1980.

References

Freeman, D.J., Kuhs, T.M., Knappen, L., & Porter, A.C. A closer look at standardized tests. The Arithmetic Teacher (in press).

Porter, A.C., Schmidt, W.H., Floden, R.E., & Freeman, D.J. Practical significance in program evaluation. American Education Research Journal, 1978, 15, 529-539.

Walker, D.F., & Schaffarzick, J. Comparing curricula. Review of Educational Research, 1974, 44(1), 83-111.